Sequence Listing

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Presta, Leonard G.
Shelton, David L.
Urfer, Roman

(ii) TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors

(iii) NUMBER OF SEQUENCES: 41

(iv) CORRESPONDENCE_ADDRESS:
(A) ADDRESSEE: Knobbe, Martens, Olson & Bear
(B) STREET: 620 Newport Center Drive 16th Floor
(C) CITY: Newport Beach
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 92660

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WinPatin (Genentech)

(vi) CURRENT APPLICATION DATA
(A) APPLICATION NUMBER: 09/724,524
(B) FILING DATE: 27-NOV-2000
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 09/156,923
(B) FILING DATE: 18-SEP-1998
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/359,705
(B) FILING DATE: 20-DEC-1994
(C) CLASSIFICATION:

(viii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/286846
(B) FILING DATE: 10-AUG-1994

(viii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/215139
(B) FILING DATE: 18-MAR-1994

(ix) ATTORNEY/AGENT INFORMATION:
(A) NAME: Dreger, Ginger
(B) REGISTRATION NUMBER: 33,055
(C) REFERENCE/DOCKET NUMBER: GENENT.33CP2C2

(x) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 949/760-0404
(B) TELEFAX: 949/760-9502

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3194 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGAAGGTTTA AAGAAGAAGC CGCAAAGCGC AGGGAAGGCC TCCCGGCACG 50

GGTGGGGGAA AGCGGCCGGT GCAGCGCGGG GACAGGCACT CGGGCTGGCA 100

CTGGCTGCTA GGGATGTCGT CCTGGATAAG GTGGCATGGA CCCGCCATGG 150

CGCGGCTCTG GGGCTTCTGC TGGCTGGTTG TGGGCTTCTG GAGGGCCGCT 200

TTCGCCTGTC CCACGTCCTG CAAATGCAGT GCCTCTCGGA TCTGGTGCAG 250

CGACCCTTCT CCTGGCATCG TGGCATTTCC GAGATTGGAG CCTAACAGTG 300

TAGATCCTGA GAACATCACC GAAATTTTCA TCGCAAACCA GAAAAGGTTA 350

GAAATCATCA ACGAAGATGA TGTTGAAGCT TATGTGGGAC TGAGAAATCT 400

GACAATTGTG GATTCTGGAT TAAAATTTGT GGCTCATAAA GCATTTCTGA 450

AAAACAGCAA CCTGCAGCAC ATCAATTTTA CCCGAAACAA ACTGACGAGT 500

TTGTCTAGGA AACATTTCCG TCACCTTGAC TTGTCTGAAC TGATCCTGGT 550

GGGCAATCCA TTTACATGCT CCTGTGACAT TATGTGGATC AAGACTCTCC 600

AAGAGGCTAA ATCCAGTCCA GACACTCAGG ATTTGTACTG CCTGAATGAA 650

AGCAGCAAGA ATATTCCCCT GGCAAACCTG CAGATACCCA ATTGTGGTTT 700

GCCATCTGCA AATCTGGCCG CACCTAACCT CACTGTGGAG GAAGGAAAGT 750

CTATCACATT ATCCTGTAGT GTGGCAGGTG ATCCGGTTCC TAATATGTAT 800

TGGGATGTTG GTAACCTGGT TTCCAAACAT ATGAATGAAA CAAGCCACAC 850

ACAGGGCTCC TTAAGGATAA CTAACATTTC ATCCGATGAC AGTGGGAAGC 900

AGATCTCTTG TGTGGCGGAA AATCTTGTAG GAGAAGATCA AGATTCTGTC 950

AACCTCACTG TGCATTTTGC ACCAACTATC ACATTTCTCG AATCTCCAAC 1000

CTCAGACCAC CACTGGTGCA TTCCATTCAC TGTGAAAGGC AACCCAAAAC 1050

CAGCGCTTCA GTGGTTCTAT AACGGGGCAA TATTGAATGA GTCCAAATAC 1100

ATCTGTACTA AAATACATGT TACCAATCAC ACGGAGTACC ACGGCTGCCT 1150

CCAGCTGGAT AATCCCACTC ACATGAACAA TGGGGACTAC ACTCTAATAG 1200

CCAAGAATGA GTATGGGAAG GATGAGAAAC AGATTTCTGC TCACTTCATG 1250

GGCTGGCCTG GAATTGACGA TGGTGCAAAC CCAAATTATC CTGATGTAAT 1300

TTATGAAGAT TATGGAACTG CAGCGAATGA CATCGGGGAC ACCACGAACA 1350

GAAGTAATGA AATCCCTTCC ACAGACGTCA CTGATAAAAC CGGTCGGGAA 1400

CATCTCTCGG TCTATGCTGT GGTGGTGATT GCGTCTGTGG TGGGATTTTG 1450

CCTTTTGGTA ATGCTGTTTC TGCTTAAGTT GGCAAGACAC TCCAAGTTTG 1500

GCATGAAAGG CCCAGCCTCC GTTATCAGCA ATGATGATGA CTCTGCCAGC 1550

CCACTCCATC ACATCTCCAA TGGGAGTAAC ACTCCATCTT CTTCGGAAGG 1600

TGGCCCAGAT GCTGTCATTA TTGGAATGAC CAAGATCCCT GTCATTGAAA 1650

ATCCCCAGTA CTTTGGCATC ACCAACAGTC AGCTCAAGCC AGACACATTT 1700

GTTCAGCACA TCAAGCGACA TAACATTGTT CTGAAAAGGG AGCTAGGCGA 1750

AGGAGCCTTT GGAAAAGTGT TCCTAGCTGA ATGCTATAAC CTCTGTCCTG 1800

AGCAGGACAA GATCTTGGTG GCAGTGAAGA CCCTGAAGGA TGCCAGTGAC 1850

AATGCACGCA AGGACTTCCA CCGTGAGGCC GAGCTCCTGA CCAACCTCCA 1900

GCATGAGCAC ATCGTCAAGT TCTATGGCGT CTGCGTGGAG GGCGACCCCC 1950

TCATCATGGT CTTTGAGTAC ATGAAGCATG GGGACCTCAA CAAGTTCCTC 2000

AGGGCACACG GCCCTGATGC CGTGCTGATG GCTGAGGGCA ACCCGCCCAC 2050

GGAACTGACG CAGTCGCAGA TGCTGCATAT AGCCCAGCAG ATCGCCGCGG 2100

GCATGGTCTA CCTGGCGTCC CAGCACTTCG TGCACCGCGA TTTGGCCACC 2150

AGGAACTGCC TGGTCGGGGA GAACTTGCTG GTGAAAATCG GGGACTTTGG 2200

GATGTCCCGG GACGTGTACA GCACTGACTA CTACAGGGTC GGTGGCCACA 2250

CAATGCTGCC CATTCGCTGG ATGCCTCCAG AGAGCATCAT GTACAGGAAA 2300

TTCACGACGG AAAGCGACGT CTGGAGCCTG GGGGTCGTGT TGTGGGAGAT 2350

TTTCACCTAT GGCAAACAGC CCTGGTACCA GCTGTCAAAC AATGAGGTGA 2400

TAGAGTGTAT CACTCAGGGC CGAGTCCTGC AGCGACCCCG CACGTGCCCC 2450

CAGGAGGTGT ATGAGCTGAT GCTGGGGTGC TGGCAGCGAG AGCCCCACAT 2500

GAGGAAGAAC ATCAAGGGCA TCCATACCCT CCTTCAGAAC TTGGCCAAGG 2550

CATCTCCGGT CTACCTGGAC ATTCTAGGCT AGGGCCCTTT TCCCCAGACC 2600

GATCCTTCCC AACGTACTCC TCAGACGGGC TGAGAGGATG AACATCTTTT 2650

AACTGCCGCT GGAGGCCACC AAGCTGCTCT CCTTCACTCT GACAGTATTA 2700

ACATCAAAGA CTCCGAGAAG CTCTCGAGGG AAGCAGTGTG TACTTCTTCA 2750

TCCATAGACA CAGTATTGAC TTCTTTTTGG CATTATCTCT TTCTCTCTTT 2800

CCATCTCCCT TGGTTGTTCC TTTTTCTTTT TTTAAATTTT CTTTTTCTTC 2850

TTTTTTTTCG TCTTCCCTGC TTCACGATTC TTACCCTTTC TTTTGAATCA 2900

ATCTGGCTTC TGCATTACTA TTAACTCTGC ATAGACAAAG GCCTTAACAA 2950

ACGTAATTTG TTATATCAGC AGACACTCCA GTTTGCCCAC CACAACTAAC 3000

AATGCCTTGT TGTATTCCTG CCTTTGATGT GGATGAAAAA AAGGGAAAAC 3050

AAATATTTCA CTTAAACTTT GTCACTTCTG CTGTACAGAT ATCGAGAGTT 3100

TCTATGGATT CACTTCTATT TATTTATTAT TATTACTGTT CTTATTGTTT 3150

TTGGATGGCT TAAGCCTGTG TATAAAAAAA AAAAAAAATC TAGA 3194

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 822 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Ser Trp Ile Arg Trp His Gly Pro Ala Met Ala Arg Leu
1 5 10 15

Trp Gly Phe Cys Trp Leu Val Val Gly Phe Trp Arg Ala Ala Phe
20 25 30

Ala Cys Pro Thr Ser Cys Lys Cys Ser Ala Ser Arg Ile Trp Cys
35 40 45

Ser Asp Pro Ser Pro Gly Ile Val Ala Phe Pro Arg Leu Glu Pro
50 55 60

Asn Ser Val Asp Pro Glu Asn Ile Thr Glu Ile Phe Ile Ala Asn
65 70 75

Gln Lys Arg Leu Glu Ile Ile Asn Glu Asp Asp Val Glu Ala Tyr
80 85 90

Val Gly Leu Arg Asn Leu Thr Ile Val Asp Ser Gly Leu Lys Phe

95 100 105

Val Ala His Lys Ala Phe Leu Lys Asn Ser Asn Leu Gln His Ile
110 115 120

Asn Phe Thr Arg Asn Lys Leu Thr Ser Leu Ser Arg Lys His Phe
125 130 135

Arg His Leu Asp Leu Ser Glu Leu Ile Leu Val Gly Asn Pro Phe
140 145 150

Thr Cys Ser Cys Asp Ile Met Trp Ile Lys Thr Leu Gln Glu Ala
155 160 165

Lys Ser Ser Pro Asp Thr Gln Asp Leu Tyr Cys Leu Asn Glu Ser
170 175 180

Ser Lys Asn Ile Pro Leu Ala Asn Leu Gln Ile Pro Asn Cys Gly
185 190 195

Leu Pro Ser Ala Asn Leu Ala Ala Pro Asn Leu Thr Val Glu Glu
200 205 210

Gly Lys Ser Ile Thr Leu Ser Cys Ser Val Ala Gly Asp Pro Val
215 220 225

Pro Asn Met Tyr Trp Asp Val Gly Asn Leu Val Ser Lys His Met
230 235 240

Asn Glu Thr Ser His Thr Gln Gly Ser Leu Arg Ile Thr Asn Ile
245 250 255

Ser Ser Asp Asp Ser Gly Lys Gln Ile Ser Cys Val Ala Glu Asn
260 265 270

Leu Val Gly Glu Asp Gln Asp Ser Val Asn Leu Thr Val His Phe
275 280 285

Ala Pro Thr Ile Thr Phe Leu Glu Ser Pro Thr Ser Asp His His
290 295 300

Trp Cys Ile Pro Phe Thr Val Lys Gly Asn Pro Lys Pro Ala Leu
305 310 315

Gln Trp Phe Tyr Asn Gly Ala Ile Leu Asn Glu Ser Lys Tyr Ile
320 325 330

Cys Thr Lys Ile His Val Thr Asn His Thr Glu Tyr His Gly Cys
335 340 345

Leu Gln Leu Asp Asn Pro Thr His Met Asn Asn Gly Asp Tyr Thr
350 355 360

Leu Ile Ala Lys Asn Glu Tyr Gly Lys Asp Glu Lys Gln Ile Ser
365 370 375

Ala His Phe Met Gly Trp Pro Gly Ile Asp Asp Gly Ala Asn Pro

380 385 390

Asn Tyr Pro Asp Val Ile Tyr Glu Asp Tyr Gly Thr Ala Ala Asn
395 400 405

Asp Ile Gly Asp Thr Thr Asn Arg Ser Asn Glu Ile Pro Ser Thr
410 415 420

Asp Val Thr Asp Lys Thr Gly Arg Glu His Leu Ser Val Tyr Ala
425 430 435

Val Val Val Ile Ala Ser Val Val Gly Phe Cys Leu Leu Val Met
440 445 450

Leu Phe Leu Leu Lys Leu Ala Arg His Ser Lys Phe Gly Met Lys
455 460 465

Gly Pro Ala Ser Val Ile Ser Asn Asp Asp Asp Ser Ala Ser Pro
470 475 480

Leu His His Ile Ser Asn Gly Ser Asn Thr Pro Ser Ser Ser Glu
485 490 495

Gly Gly Pro Asp Ala Val Ile Ile Gly Met Thr Lys Ile Pro Val
500 505 510

Ile Glu Asn Pro Gln Tyr Phe Gly Ile Thr Asn Ser Gln Leu Lys
515 520 525

Pro Asp Thr Phe Val Gln His Ile Lys Arg His Asn Ile Val Leu
530 535 540

Lys Arg Glu Leu Gly Glu Gly Ala Phe Gly Lys Val Phe Leu Ala
545 550 555

Glu Cys Tyr Asn Leu Cys Pro Glu Gln Asp Lys Ile Leu Val Ala
560 565 570

Val Lys Thr Leu Lys Asp Ala Ser Asp Asn Ala Arg Lys Asp Phe
575 580 585

His Arg Glu Ala Glu Leu Leu Thr Asn Leu Gln His Glu His Ile
590 595 600

Val Lys Phe Tyr Gly Val Cys Val Glu Gly Asp Pro Leu Ile Met
605 610 615

Val Phe Glu Tyr Met Lys His Gly Asp Leu Asn Lys Phe Leu Arg
620 625 630

Ala His Gly Pro Asp Ala Val Leu Met Ala Glu Gly Asn Pro Pro
635 640 645

Thr Glu Leu Thr Gln Ser Gln Met Leu His Ile Ala Gln Gln Ile
650 655 660

Ala Ala Gly Met Val Tyr Leu Ala Ser Gln His Phe Val His Arg

```
                665                 670                 675

Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu Asn Leu Leu Val
                680                 685                 690

Lys Ile Gly Asp Phe Gly Met Ser Arg Asp Val Tyr Ser Thr Asp
                695                 700                 705

Tyr Tyr Arg Val Gly Gly His Thr Met Leu Pro Ile Arg Trp Met
                710                 715                 720

Pro Pro Glu Ser Ile Met Tyr Arg Lys Phe Thr Thr Glu Ser Asp
                725                 730                 735

Val Trp Ser Leu Gly Val Val Leu Trp Glu Ile Phe Thr Tyr Gly
                740                 745                 750

Lys Gln Pro Trp Tyr Gln Leu Ser Asn Asn Glu Val Ile Glu Cys
                755                 760                 765

Ile Thr Gln Gly Arg Val Leu Gln Arg Pro Arg Thr Cys Pro Gln
                770                 775                 780

Glu Val Tyr Glu Leu Met Leu Gly Cys Trp Gln Arg Glu Pro His
                785                 790                 795

Met Arg Lys Asn Ile Lys Gly Ile His Thr Leu Leu Gln Asn Leu
                800                 805                 810

Ala Lys Ala Ser Pro Val Tyr Leu Asp Ile Leu Gly
                815                 820     822
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1870 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GGAAGGTTTA AAGAAGAAGC CGCAAAGCGC AGGGAAGGCC TCCCGGCACG 50

GGTGGGGGAA AGCGGCCGGT GCAGCGCGGG GACAGGCACT CGGGCTGGCA 100

CTGGCTGCTA GGGATGTCGT CCTGGATAAG GTGGGATGGA CCCGCCATGG 150

CGCGGCTCTG GGGCTTCTGC TGGCTGGTTG TGGGCTTCTG GAGGGCCGCT 200

TTCGCCTGTC CCACGTCCTG CAAATGCAGT GCCTCTCGGA TCTGGTGCAG 250

CGACCCTTCT CCTGGCATCG TGGCATTTCC GAGATTGGAG CCTAACAGTG 300

TAGATCCTGA GAACATCACC GAAATTTTCA TCGCAAACCA GAAAAGGTTA 350
```

GAAATCATCA ACGAAGATGA TGTTGAAGCT TATGTGGGAC TGAGAAATCT 400

GACAATTGTG GATTCTGGAT TAAAATTTGT GGCTCATAAA GCATTTCTGA 450

AAAACAGCAA CCTGCAGCAC ATCAATTTTA CCCGAAACAA ACTGACGAGT 500

TTGTCTAGGA AACATTTCCG TCACCTTGAC TTGTCTGAAC TGATCCTGGT 550

GGGCAATCCA TTTACATGCT CCTGTGACAT TATGTGGATC AAGACTCTCC 600

AAGAGGCTAA ATCCAGTCCA GACACTCAGG ATTTGTACTG CCTGAATGAA 650

AGCAGCAAGA ATATTCCCCT GGCAAACCTG CAGATACCCA ATTGTGGTTT 700

GCCATCTGCA AATCTGGCCG CACCTAACCT CACTGTGGAG GAAGGAAAGT 750

CTATCACATT ATCCTGTAGT GTGGCAGGTG ATCCGGTTCC TAATATGTAT 800

TGGGATGTTG GTAACCTGGT TTCCAAACAT ATGAATGAAA CAAGCCACAC 850

ACAGGGCTCC TTAAGGATAA CTAACATTTC ATCCGATGAC AGTGGGAAGC 900

AGATCTCTTG TGTGGCGGAA AATCTTGTAG GAGAAGATCA AGATTCTGTC 950

AACCTCACTG TGCATTTTGC ACCAACTATC ACATTTCTCG AATCTCCAAC 1000

CTCAGACCAC CACTGGTGCA TTCCATTCAC TGTGAAAGGC AACCCAAAAC 1050

CAGCGCTTCA GTGGTTCTAT AACGGGGCAA TATTGAATGA GTCCAAATAC 1100

ATCTGTACTA AAATACATGT TACCAATCAC ACGGAGTACC ACGGCTGCCT 1150

CCAGCTGGAT AATCCCACTC ACATGAACAA TGGGGACTAC ACTCTAATAG 1200

CCAAGAATGA GTATGGGAAG GATGAGAAAC AGATTTCTGC TCACTTCATG 1250

GGCTGGCCTG GAATTGACGA TGGTGCAAAC CCAAATTATC CTGATGTAAT 1300

TTATGAAGAT TATGGAACTG CAGCGAATGA CATCGGGGAC ACCACGAACA 1350

GAAGTAATGA AATCCCTTCC ACAGACGTCA CTGATAAAAC CGGTCGGGAA 1400

CATCTCTCGG TCTATGCTGT GGTGGTGATT GCGTCTGTGG TGGGATTTTG 1450

CCTTTTGGTA ATGCTGTTTC TGCTTAAGTT GGCAAGACAC TCCAAGTTTG 1500

GCATGAAAGG TTTTGTTTTG TTTCATAAGA TCCCACTGGA TGGGTAGCTG 1550

AAATAAAGGA AAAGACAGAG AAAGGGGCTG TGGTGCTTGT TGGTTGATGC 1600

TGCCATGTAA GCTGGACTCC TGGGACTGCT GTTGGCTTAT CCCGGGAAGT 1650

GCTGCTTATC TGGGGTTTTC TGGTAGATGT GGGCGGTGTT TGGAGGCTGT 1700

ACTATATGAA GCCTGCATAT ACTGTGAGCT GTGATTGGGG AACACCAATG 1750

CAGAGGTAAC TCTCAGGCAG CTAAGCAGCA CCTCAAGAAA ACATGTTAAA 1800

TTAATGCTTC TCTTCTTACA GTAGTTCAAA TACAAAACTG AAATGAAATC 1850

CCATTGGATT GTACTTCTCT 1870

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 477 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Ser Trp Ile Arg Trp His Gly Pro Ala Met Ala Arg Leu
1 5 10 15

Trp Gly Phe Cys Trp Leu Val Val Gly Phe Trp Arg Ala Ala Phe
20 25 30

Ala Cys Pro Thr Ser Cys Lys Cys Ser Ala Ser Arg Ile Trp Cys
35 40 45

Ser Asp Pro Ser Pro Gly Ile Val Ala Phe Pro Arg Leu Glu Pro
50 55 60

Asn Ser Val Asp Pro Glu Asn Ile Thr Glu Ile Phe Ile Ala Asn
65 70 75

Gln Lys Arg Leu Glu Ile Ile Asn Glu Asp Asp Val Glu Ala Tyr
80 85 90

Val Gly Leu Arg Asn Leu Thr Ile Val Asp Ser Gly Leu Lys Phe
95 100 105

Val Ala His Lys Ala Phe Leu Lys Asn Ser Asn Leu Gln His Ile
110 115 120

Asn Phe Thr Arg Asn Lys Leu Thr Ser Leu Ser Arg Lys His Phe
125 130 135

Arg His Leu Asp Leu Ser Glu Leu Ile Leu Val Gly Asn Pro Phe
140 145 150

Thr Cys Ser Cys Asp Ile Met Trp Ile Lys Thr Leu Gln Glu Ala
155 160 165

Lys Ser Ser Pro Asp Thr Gln Asp Leu Tyr Cys Leu Asn Glu Ser
170 175 180

Ser Lys Asn Ile Pro Leu Ala Asn Leu Gln Ile Pro Asn Cys Gly
185 190 195

Leu Pro Ser Ala Asn Leu Ala Ala Pro Asn Leu Thr Val Glu Glu
200 205 210

Gly Lys Ser Ile Thr Leu Ser Cys Ser Val Ala Gly Asp Pro Val

215 220 225

Pro Asn Met Tyr Trp Asp Val Gly Asn Leu Val Ser Lys His Met
230 235 240

Asn Glu Thr Ser His Thr Gln Gly Ser Leu Arg Ile Thr Asn Ile
245 250 255

Ser Ser Asp Asp Ser Gly Lys Gln Ile Ser Cys Val Ala Glu Asn
260 265 270

Leu Val Gly Glu Asp Gln Asp Ser Val Asn Leu Thr Val His Phe
275 280 285

Ala Pro Thr Ile Thr Phe Leu Glu Ser Pro Thr Ser Asp His His
290 295 300

Trp Cys Ile Pro Phe Thr Val Lys Gly Asn Pro Lys Pro Ala Leu
305 310 315

Gln Trp Phe Tyr Asn Gly Ala Ile Leu Asn Glu Ser Lys Tyr Ile
320 325 330

Cys Thr Lys Ile His Val Thr Asn His Thr Glu Tyr His Gly Cys
335 340 345

Leu Gln Leu Asp Asn Pro Thr His Met Asn Asn Gly Asp Tyr Thr
350 355 360

Leu Ile Ala Lys Asn Glu Tyr Gly Lys Asp Glu Lys Gln Ile Ser
365 370 375

Ala His Phe Met Gly Trp Pro Gly Ile Asp Asp Gly Ala Asn Pro
380 385 390

Asn Tyr Pro Asp Val Ile Tyr Glu Asp Tyr Gly Thr Ala Ala Asn
395 400 405

Asp Ile Gly Asp Thr Thr Asn Arg Ser Asn Glu Ile Pro Ser Thr
410 415 420

Asp Val Thr Asp Lys Thr Gly Arg Glu His Leu Ser Val Tyr Ala
425 430 435

Val Val Val Ile Ala Ser Val Val Gly Phe Cys Leu Leu Val Met
440 445 450

Leu Phe Leu Leu Lys Leu Ala Arg His Ser Lys Phe Gly Met Lys
455 460 465

Gly Phe Val Leu Phe His Lys Ile Pro Leu Asp Gly
470 475 477

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2715 base pairs

(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGATCCGCGT CGGAGATGGA TGTCTCTCTT TGCCCAGCCA AGTGTAGTTT 50

CTGGCGGATT TTCTTGCTGG GAAGCGTCTG GCTGGACTAT GTGGGCTCCG 100

TGCTGGCTTG CCCTGCAAAT TGTGTCTGCA GCAAGACTGA GATCAATTGC 150

CGGCGGCCGG ACGATGGGAA CCTCTTCCCC CTCCTGGAAG GGCAGGATTC 200

AGGGAACAGC AATGGGAACG CCAATATCAA CATCACGGAC ATCTCAAGGA 250

ATATCACTTC CATACACATA GAGAACTGGC GCAGTCTTCA CACGCTCAAC 300

GCCGTGGACA TGGAGCTCTA CACCGGACTT CAAAAGCTGA CCATCAAGAA 350

CTCAGGACTT CGGAGCATTC AGCCCAGAGC CTTTGCCAAG AACCCCCATT 400

TGCGTTATAT AAACCTGTCA AGTAACCGGC TCACCACACT CTCGTGGCAG 450

CTCTTCCAGA CGCTGAGTCT TCGGGAATTG CAGTTGGAGC AGAACTTTTT 500

CAACTGCAGC TGTGACATCC GCTGGATGCA GCTCTGGCAG GAGCAGGGGG 550

AGGCCAAGCT CAACAGCCAG AACCTCTACT GCATCAATGC TGATGGCTCC 600

CAGCTTCCTC TCTTCCGCAT GAACATCAGT CAGTGTGACC TTCCTGAGAT 650

CAGCGTGAGC CACGTCAACC TGACCGTACG AGAGGGTGAC AATGCTGTTA 700

TCACTTGCAA TGGCTCTGGA TCACCCCTTC CTGATGTGGA CTGGATAGTC 750

ACTGGGCTGC AGTCCATCAA CACTCACCAG ACCAATCTGA ACTGGACCAA 800

TGTTCATGCC ATCAACTTGA CGCTGGTGAA TGTGACGAGT GAGGACAATG 850

GCTTCACCCT GACGTGCATT GCAGAGAACG TGGTGGGCAT GAGCAATGCC 900

AGTGTTGCCC TCACTGTCTA CTATCCCCCA CGTGTGGTGA GCCTGGAGGA 950

GCCTGAGCTG CGCCTGGAGC ACTGCATCGA GTTTGTGGTG CGTGGCAACC 1000

CCCCACCAAC GCTGCACTGG CTGCACAATG GGCAGCCTCT GCGGGAGTCC 1050

AAGATCATCC ATGTGGAATA CTACCAAGAG GGAGAGATTT CCGAGGGCTG 1100

CCTGCTCTTC AACAAGCCCA CCCACTACAA CAATGGCAAC TATACCCTCA 1150

TTGCCAAAAA CCCACTGGGC ACAGCCAACC AGACCATCAA TGGCCACTTC 1200

CTCAAGGAGC CCTTTCCAGA GAGCACGGAT AACTTTATCT TGTTTGACGA 1250

AGTGAGTCCC ACACCTCCTA TCACTGTGAC CCACAAACCA GAAGAAGACA 1300

CTTTTGGGGT ATCCATAGCA GTTGGACTTG CTGCTTTTGC CTGTGTCCTG 1350

TTGGTGGTTC TCTTCGTCAT GATCAACAAA TATGGTCGAC GGTCCAAATT 1400

TGGAATGAAG GGTCCCGTGG CTGTCATCAG TGGTGAGGAG GACTCAGCCA 1450

GCCCACTGCA CCACATCAAC CACGGCATCA CCACGCCCTC GTCACTGGAT 1500

GCCGGGCCCG ACACTGTGGT CATTGGCATG ACTCGCATCC CTGTCATTGA 1550

GAACCCCCAG TACTTCCGTC AGGGACACAA CTGCCACAAG CCGGACACGT 1600

ATGTGCAGCA CATTAAGAGG AGAGACATCG TGCTGAAGCG AGAACTGGGT 1650

GAGGGAGCCT TTGGAAAGGT CTTCCTGGCC GAGTGCTACA ACCTCAGCCC 1700

GACCAAGGAC AAGATGCTTG TGGCTGTGAA GGCCCTGAAG GATCCCACCC 1750

TGGCTGCCCG GAAGGATTTC CAGAGGGAGG CCGAGCTGCT CACCAACCTG 1800

CAGCATGAGC ACATTGTCAA GTTCTATGGA GTGTGCGGCG ATGGGGACCC 1850

CCTCATCATG GTCTTTGAAT ACATGAAGCA TGGAGACCTG AATAAGTTCC 1900

TCAGGGCCCA TGGGCCAGAT GCAATGATCC TTGTGGATGG ACAGCCACGC 1950

CAGGCCAAGG GTGAGCTGGG GCTCTCCCAA ATGCTCCACA TTGCCAGTCA 2000

GATCGCCTCG GGTATGGTGT ACCTGGCCTC CCAGCACTTT GTGCACCGAG 2050

ACCTGGCCAC CAGGAACTGC CTGGTTGGAG CGAATCTGCT AGTGAAGATT 2100

GGGGACTTCG GCATGTCCAG AGATGTCTAC AGCACGGATT ATTACAGGCT 2150

CTTTAATCCA TCTGGAAATG ATTTTTGTAT ATGGTGTGAG GTGGGAGGAC 2200

ACACCATGCT CCCCATTCGC TGGATGCCTC CTGAAAGCAT CATGTACCGG 2250

AAGTTCACTA CAGAGAGTGA TGTATGGAGC TTCGGGGTGA TCCTCTGGGA 2300

GATCTTCACC TATGGAAAGC AGCCATGGTT CCAACTCTCA AACACGGAGG 2350

TCATTGAGTG CATTACCCAA GGTCGTGTTT TGGAGCGGCC CCGAGTCTGC 2400

CCCAAAGAGG TGTACGATGT CATGCTGGGG TGCTGGCAGA GGGAACCACA 2450

GCAGCGGTTG AACATCAAGG AGATCTACAA AATCCTCCAT GCTTTGGGGA 2500

AGGCCACCCC AATCTACCTG GACATTCTTG GCTAGTGGTG GCTGGTGGTC 2550

ATGAATTCAT ACTCTGTTGC CTCCTCTCTC CCTGCCTCAC ATCTCCCTTC 2600

CACCTCACAA CTCCTTCCAT CCTTGACTGA AGCGAACATC TTCATATAAA 2650

CTCAAGTGCC TGCTACACAT ACAACACTGA AAAAAGGAAA AAAAAAGAAA 2700

AAAAAAAAAA ACCGC 2715

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 839 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Asp Val Ser Leu Cys Pro Ala Lys Cys Ser Phe Trp Arg Ile
 1               5                  10                  15

Phe Leu Leu Gly Ser Val Trp Leu Asp Tyr Val Gly Ser Val Leu
            20                  25                  30

Ala Cys Pro Ala Asn Cys Val Cys Ser Lys Thr Glu Ile Asn Cys
        35                  40                  45

Arg Arg Pro Asp Asp Gly Asn Leu Phe Pro Leu Leu Glu Gly Gln
    50                  55                  60

Asp Ser Gly Asn Ser Asn Gly Asn Ala Asn Ile Asn Ile Thr Asp
65                  70                  75

Ile Ser Arg Asn Ile Thr Ser Ile His Ile Glu Asn Trp Arg Ser
            80                  85                  90

Leu His Thr Leu Asn Ala Val Asp Met Glu Leu Tyr Thr Gly Leu
        95                  100                 105

Gln Lys Leu Thr Ile Lys Asn Ser Gly Leu Arg Ser Ile Gln Pro
    110                 115                 120

Arg Ala Phe Ala Lys Asn Pro His Leu Arg Tyr Ile Asn Leu Ser
125                 130                 135

Ser Asn Arg Leu Thr Thr Leu Ser Trp Gln Leu Phe Gln Thr Leu
            140                 145                 150

Ser Leu Arg Glu Leu Gln Leu Glu Gln Asn Phe Phe Asn Cys Ser
        155                 160                 165

Cys Asp Ile Arg Trp Met Gln Leu Trp Gln Glu Gln Gly Glu Ala
    170                 175                 180

Lys Leu Asn Ser Gln Asn Leu Tyr Cys Ile Asn Ala Asp Gly Ser
185                 190                 195

Gln Leu Pro Leu Phe Arg Met Asn Ile Ser Gln Cys Asp Leu Pro
            200                 205                 210

Glu Ile Ser Val Ser His Val Asn Leu Thr Val Arg Glu Gly Asp
        215                 220                 225
```

Asn Ala Val Ile Thr Cys Asn Gly Ser Gly Ser Pro Leu Pro Asp
230 235 240

Val Asp Trp Ile Val Thr Gly Leu Gln Ser Ile Asn Thr His Gln
245 250 255

Thr Asn Leu Asn Trp Thr Asn Val His Ala Ile Asn Leu Thr Leu
260 265 270

Val Asn Val Thr Ser Glu Asp Asn Gly Phe Thr Leu Thr Cys Ile
275 280 285

Ala Glu Asn Val Val Gly Met Ser Asn Ala Ser Val Ala Leu Thr
290 295 300

Val Tyr Tyr Pro Pro Arg Val Val Ser Leu Glu Glu Pro Glu Leu
305 310 315

Arg Leu Glu His Cys Ile Glu Phe Val Val Arg Gly Asn Pro Pro
320 325 330

Pro Thr Leu His Trp Leu His Asn Gly Gln Pro Leu Arg Glu Ser
335 340 345

Lys Ile Ile His Val Glu Tyr Tyr Gln Glu Gly Glu Ile Ser Glu
350 355 360

Gly Cys Leu Leu Phe Asn Lys Pro Thr His Tyr Asn Asn Gly Asn
365 370 375

Tyr Thr Leu Ile Ala Lys Asn Pro Leu Gly Thr Ala Asn Gln Thr
380 385 390

Ile Asn Gly His Phe Leu Lys Glu Pro Phe Pro Glu Ser Thr Asp
395 400 405

Asn Phe Ile Leu Phe Asp Glu Val Ser Pro Thr Pro Pro Ile Thr
410 415 420

Val Thr His Lys Pro Glu Glu Asp Thr Phe Gly Val Ser Ile Ala
425 430 435

Val Gly Leu Ala Ala Phe Ala Cys Val Leu Leu Val Val Leu Phe
440 445 450

Val Met Ile Asn Lys Tyr Gly Arg Arg Ser Lys Phe Gly Met Lys
455 460 465

Gly Pro Val Ala Val Ile Ser Gly Glu Glu Asp Ser Ala Ser Pro
470 475 480

Leu His His Ile Asn His Gly Ile Thr Thr Pro Ser Ser Leu Asp
485 490 495

Ala Gly Pro Asp Thr Val Val Ile Gly Met Thr Arg Ile Pro Val
500 505 510

Ile Glu Asn Pro Gln Tyr Phe Arg Gln Gly His Asn Cys His Lys
515 520 525

Pro Asp Thr Tyr Val Gln His Ile Lys Arg Arg Asp Ile Val Leu
530 535 540

Lys Arg Glu Leu Gly Glu Gly Ala Phe Gly Lys Val Phe Leu Ala
545 550 555

Glu Cys Tyr Asn Leu Ser Pro Thr Lys Asp Lys Met Leu Val Ala
560 565 570

Val Lys Ala Leu Lys Asp Pro Thr Leu Ala Ala Arg Lys Asp Phe
575 580 585

Gln Arg Glu Ala Glu Leu Leu Thr Asn Leu Gln His Glu His Ile
590 595 600

Val Lys Phe Tyr Gly Val Cys Gly Asp Gly Asp Pro Leu Ile Met
605 610 615

Val Phe Glu Tyr Met Lys His Gly Asp Leu Asn Lys Phe Leu Arg
620 625 630

Ala His Gly Pro Asp Ala Met Ile Leu Val Asp Gly Gln Pro Arg
635 640 645

Gln Ala Lys Gly Glu Leu Gly Leu Ser Gln Met Leu His Ile Ala
650 655 660

Ser Gln Ile Ala Ser Gly Met Val Tyr Leu Ala Ser Gln His Phe
665 670 675

Val His Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Ala Asn
680 685 690

Leu Leu Val Lys Ile Gly Asp Phe Gly Met Ser Arg Asp Val Tyr
695 700 705

Ser Thr Asp Tyr Tyr Arg Leu Phe Asn Pro Ser Gly Asn Asp Phe
710 715 720

Cys Ile Trp Cys Glu Val Gly Gly His Thr Met Leu Pro Ile Arg
725 730 735

Trp Met Pro Pro Glu Ser Ile Met Tyr Arg Lys Phe Thr Thr Glu
740 745 750

Ser Asp Val Trp Ser Phe Gly Val Ile Leu Trp Glu Ile Phe Thr
755 760 765

Tyr Gly Lys Gln Pro Trp Phe Gln Leu Ser Asn Thr Glu Val Ile
770 775 780

Glu Cys Ile Thr Gln Gly Arg Val Leu Glu Arg Pro Arg Val Cys
785 790 795

Pro Lys Glu Val Tyr Asp Val Met Leu Gly Cys Trp Gln Arg Glu
800 805 810

Pro Gln Gln Arg Leu Asn Ile Lys Glu Ile Tyr Lys Ile Leu His
815 820 825

Ala Leu Gly Lys Ala Thr Pro Ile Tyr Leu Asp Ile Leu Gly
830 835 839

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1858 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGATCCGCGT CGGAGATGGA TGTCTCTCTT TGCCCAGCCA AGTGTAGTTT 50

CTGGCGGATT TTCTTGCTGG GAAGCGTCTG GCTGGACTAT GTGGGCTCCG 100

TGCTGGCTTG CCCTGCAAAT TGTGTCTGCA GCAAGACTGA GATCAATTGC 150

CGGCGGCCGG ACGATGGGAA CCTCTTCCCC CTCCTGGAAG GGCAGGATTC 200

AGGGAACAGC AATGGGAACG CCAATATCAA CATCACGGAC ATCTCAAGGA 250

ATATCACTTC CATACACATA GAGAACTGGC GCAGTCTTCA CACGCTCAAC 300

GCCGTGGACA TGGAGCTCTA CACCGGACTT CAAAAGCTGA CCATCAAGAA 350

CTCAGGACTT CGGAGCATTC AGCCCAGAGC CTTTGCCAAG AACCCCCATT 400

TGCGTTATAT AAACCTGTCA AGTAACCGGC TCACCACACT CTCGTGGCAG 450

CTCTTCCAGA CGCTGAGTCT TCGGGAATTG CAGTTGGAGC AGAACTTTTT 500

CAACTGCAGC TGTGACATCC GCTGGATGCA GCTCTGGCAG GAGCAGGGGG 550

AGGCCAAGCT CAACAGCCAG AACCTCTACT GCATCAATGC TGATGGCTCC 600

CAGCTTCCTC TCTTCCGCAT GAACATCAGT CAGTGTGACC TTCCTGAGAT 650

CAGCGTGAGC CACGTCAACC TGACCGTACG AGAGGGTGAC AATGCTGTTA 700

TCACTTGCAA TGGCTCTGGA TCACCCCTTC CTGATGTGGA CTGGATAGTC 750

ACTGGGCTGC AGTCCATCAA CACTCACCAG ACCAATCTGA ACTGGACCAA 800

TGTTCATGCC ATCAACTTGA CGCTGGTGAA TGTGACGAGT GAGGACAATG 850

GCTTCACCCT GACGTGCATT GCAGAGAACG TGGTGGGCAT GAGCAATGCC 900

AGTGTTGCCC TCACTGTCTA CTATCCCCCA CGTGTGGTGA GCCTGGAGGA 950

GCCTGAGCTG CGCCTGGAGC ACTGCATCGA GTTTGTGGTG CGTGGCAACC 1000

CCCCACCAAC GCTGCACTGG CTGCACAATG GGCAGCCTCT GCGGGAGTCC 1050

AAGATCATCC ATGTGGAATA CTACCAAGAG GGAGAGATTT CCGAGGGCTG 1100

CCTGCTCTTC AACAAGCCCA CCCACTACAA CAATGGCAAC TATACCCTCA 1150

TTGCCAAAAA CCCACTGGGC ACAGCCAACC AGACCATCAA TGGCCACTTC 1200

CTCAAGGAGC CCTTTCCAGA GAGCACGGAT AACTTTATCT TGTTTGACGA 1250

AGTGAGTCCC ACACCTCCTA TCACTGTGAC CCACAAACCA GAAGAAGACA 1300

CTTTTGGGGT ATCCATAGCA GTTGGACTTG CTGCTTTTGC CTGTGTCCTG 1350

TTGGTGGTTC TCTTCGTCAT GATCAACAAA TATGGTCGAC GGTCCAAATT 1400

TGGAATGAAG GGTCCCGTGG CTGTCATCAG TGGTGAGGAG GACTCAGCCA 1450

GCCCACTGCA CCACATCAAC CACGGCATCA CCACGCCCTC GTCACTGGAT 1500

GCCGGGCCCG ACACTGTGGT CATTGGCATG ACTCGCATCC CTGTCATTGA 1550

GAACCCCCAG TACTTCCGTC AGGGACACAA CTGCCACAAG CCGGACACGT 1600

GGGTCTTTTC AAACATAGAC AATCATGGGA TATTAAACTT GAAGGACAAT 1650

AGAGATCATC TAGTCCCATC AACTCACTAT ATATATGAGG AACCTGAGGT 1700

CCAGAGTGGG GAAGTGTCTT ACCCAAGGTC ACATGGTTTC AGAGAAATTA 1750

TGTTGAATCC AATAAGCCTT CCCGGACATT CCAAGCCTCT TAACCATGGC 1800

ATCTATGTTG AGGATGTCAA TGTTTATTTC AGCAAAGGAC GTCATGGCCT 1850

TTAAAAAC 1858

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 612 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Asp Val Ser Leu Cys Pro Ala Lys Cys Ser Phe Trp Arg Ile
1 5 10 15

Phe Leu Leu Gly Ser Val Trp Leu Asp Tyr Val Gly Ser Val Leu
20 25 30

Ala Cys Pro Ala Asn Cys Val Cys Ser Lys Thr Glu Ile Asn Cys
35 40 45

Arg Arg Pro Asp Asp Gly Asn Leu Phe Pro Leu Leu Glu Gly Gln
50 55 60

Asp Ser Gly Asn Ser Asn Gly Asn Ala Asn Ile Asn Ile Thr Asp
65 70 75

Ile Ser Arg Asn Ile Thr Ser Ile His Ile Glu Asn Trp Arg Ser
80 85 90

Leu His Thr Leu Asn Ala Val Asp Met Glu Leu Tyr Thr Gly Leu
95 100 105

Gln Lys Leu Thr Ile Lys Asn Ser Gly Leu Arg Ser Ile Gln Pro
110 115 120

Arg Ala Phe Ala Lys Asn Pro His Leu Arg Tyr Ile Asn Leu Ser
125 130 135

Ser Asn Arg Leu Thr Thr Leu Ser Trp Gln Leu Phe Gln Thr Leu
140 145 150

Ser Leu Arg Glu Leu Gln Leu Glu Gln Asn Phe Phe Asn Cys Ser
155 160 165

Cys Asp Ile Arg Trp Met Gln Leu Trp Gln Glu Gln Gly Glu Ala
170 175 180

Lys Leu Asn Ser Gln Asn Leu Tyr Cys Ile Asn Ala Asp Gly Ser
185 190 195

Gln Leu Pro Leu Phe Arg Met Asn Ile Ser Gln Cys Asp Leu Pro
200 205 210

Glu Ile Ser Val Ser His Val Asn Leu Thr Val Arg Glu Gly Asp
215 220 225

Asn Ala Val Ile Thr Cys Asn Gly Ser Gly Ser Pro Leu Pro Asp
230 235 240

Val Asp Trp Ile Val Thr Gly Leu Gln Ser Ile Asn Thr His Gln
245 250 255

Thr Asn Leu Asn Trp Thr Asn Val His Ala Ile Asn Leu Thr Leu
260 265 270

Val Asn Val Thr Ser Glu Asp Asn Gly Phe Thr Leu Thr Cys Ile
275 280 285

Ala Glu Asn Val Val Gly Met Ser Asn Ala Ser Val Ala Leu Thr
290 295 300

Val Tyr Tyr Pro Pro Arg Val Val Ser Leu Glu Glu Pro Glu Leu
305 310 315

Arg Leu Glu His Cys Ile Glu Phe Val Val Arg Gly Asn Pro Pro
320 325 330

```
Pro Thr Leu His Trp Leu His Asn Gly Gln Pro Leu Arg Glu Ser
                335                 340                 345

Lys Ile Ile His Val Glu Tyr Tyr Gln Glu Gly Glu Ile Ser Glu
                350                 355                 360

Gly Cys Leu Leu Phe Asn Lys Pro Thr His Tyr Asn Asn Gly Asn
                365                 370                 375

Tyr Thr Leu Ile Ala Lys Asn Pro Leu Gly Thr Ala Asn Gln Thr
                380                 385                 390

Ile Asn Gly His Phe Leu Lys Glu Pro Phe Pro Glu Ser Thr Asp
                395                 400                 405

Asn Phe Ile Leu Phe Asp Glu Val Ser Pro Thr Pro Pro Ile Thr
                410                 415                 420

Val Thr His Lys Pro Glu Glu Asp Thr Phe Gly Val Ser Ile Ala
                425                 430                 435

Val Gly Leu Ala Ala Phe Ala Cys Val Leu Leu Val Val Leu Phe
                440                 445                 450

Val Met Ile Asn Lys Tyr Gly Arg Arg Ser Lys Phe Gly Met Lys
                455                 460                 465

Gly Pro Val Ala Val Ile Ser Gly Glu Glu Asp Ser Ala Ser Pro
                470                 475                 480

Leu His His Ile Asn His Gly Ile Thr Thr Pro Ser Ser Leu Asp
                485                 490                 495

Ala Gly Pro Asp Thr Val Val Ile Gly Met Thr Arg Ile Pro Val
                500                 505                 510

Ile Glu Asn Pro Gln Tyr Phe Arg Gln Gly His Asn Cys His Lys
                515                 520                 525

Pro Asp Thr Trp Val Phe Ser Asn Ile Asp Asn His Gly Ile Leu
                530                 535                 540

Asn Leu Lys Asp Asn Arg Asp His Leu Val Pro Ser Thr His Tyr
                545                 550                 555

Ile Tyr Glu Glu Pro Glu Val Gln Ser Gly Glu Val Ser Tyr Pro
                560                 565                 570

Arg Ser His Gly Phe Arg Glu Ile Met Leu Asn Pro Ile Ser Leu
                575                 580                 585

Pro Gly His Ser Lys Pro Leu Asn His Gly Ile Tyr Val Glu Asp
                590                 595                 600

Val Asn Val Tyr Phe Ser Lys Gly Arg His Gly Phe
                605                 610     612
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 790 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Leu Arg Gly Gly Arg Arg Gly Gln Leu Gly Trp His Ser Trp
 1               5                  10                  15

Ala Ala Gly Pro Gly Ser Leu Leu Ala Trp Leu Ile Leu Ala Ser
            20                  25                  30

Ala Gly Ala Ala Pro Cys Pro Asp Ala Cys Cys Pro His Gly Ser
            35                  40                  45

Ser Gly Leu Arg Cys Thr Arg Asp Gly Ala Leu Asp Ser Leu His
            50                  55                  60

His Leu Pro Gly Ala Glu Asn Leu Thr Glu Leu Tyr Ile Glu Asn
            65                  70                  75

Gln Gln His Leu Gln His Leu Glu Leu Arg Asp Leu Arg Gly Leu
            80                  85                  90

Gly Glu Leu Arg Asn Leu Thr Ile Val Lys Ser Gly Leu Arg Phe
            95                 100                 105

Val Ala Pro Asp Ala Phe His Phe Thr Pro Arg Leu Ser Arg Leu
           110                 115                 120

Asn Leu Ser Phe Asn Ala Leu Glu Ser Leu Ser Trp Lys Thr Val
           125                 130                 135

Gln Gly Leu Ser Leu Gln Glu Leu Val Leu Ser Gly Asn Pro Leu
           140                 145                 150

His Cys Ser Cys Ala Leu Arg Trp Leu Gln Arg Trp Glu Glu Glu
           155                 160                 165

Gly Leu Gly Gly Val Pro Glu Gln Lys Leu Gln Cys His Gly Gln
           170                 175                 180

Gly Pro Leu Ala His Met Pro Asn Ala Ser Cys Gly Val Pro Thr
           185                 190                 195

Leu Lys Val Gln Val Pro Asn Ala Ser Val Asp Val Gly Asp Asp
           200                 205                 210

Val Leu Leu Arg Cys Gln Val Glu Gly Arg Gly Leu Glu Gln Ala
           215                 220                 225

Gly Trp Ile Leu Thr Glu Leu Glu Gln Ser Ala Thr Val Met Lys
           230                 235                 240
```

Ser Gly Gly Leu Pro Ser Leu Gly Leu Thr Leu Ala Asn Val Thr
245 250 255

Ser Asp Leu Asn Arg Lys Asn Leu Thr Cys Trp Ala Glu Asn Asp
260 265 270

Val Gly Arg Ala Glu Val Ser Val Gln Val Asn Val Ser Phe Pro
275 280 285

Ala Ser Val Gln Leu His Thr Ala Val Glu Met His His Trp Cys
290 295 300

Ile Pro Phe Ser Val Asp Gly Gln Pro Ala Pro Ser Leu Arg Trp
305 310 315

Leu Phe Asn Gly Ser Val Leu Asn Glu Thr Ser Phe Ile Phe Thr
320 325 330

Glu Phe Leu Glu Pro Ala Ala Asn Glu Thr Val Arg His Gly Cys
335 340 345

Leu Arg Leu Asn Gln Pro Thr His Val Asn Asn Gly Asn Tyr Thr
350 355 360

Leu Leu Ala Ala Asn Pro Phe Gly Gln Ala Ser Ala Ser Ile Met
365 370 375

Ala Ala Phe Met Asp Asn Pro Phe Glu Phe Asn Pro Glu Asp Pro
380 385 390

Ile Pro Asp Thr Asn Ser Thr Ser Gly Asp Pro Val Glu Lys Lys
395 400 405

Asp Glu Thr Pro Phe Gly Val Ser Val Ala Val Gly Leu Ala Val
410 415 420

Phe Ala Cys Leu Phe Leu Ser Thr Leu Leu Leu Val Leu Asn Lys
425 430 435

Cys Gly Arg Arg Asn Lys Phe Gly Ile Asn Arg Pro Ala Val Leu
440 445 450

Ala Pro Glu Asp Gly Leu Ala Met Ser Leu His Phe Met Thr Leu
455 460 465

Gly Gly Ser Ser Leu Ser Pro Thr Glu Gly Lys Gly Ser Gly Leu
470 475 480

Gln Gly His Ile Ile Glu Asn Pro Gln Tyr Phe Ser Asp Ala Cys
485 490 495

Val His His Ile Lys Arg Arg Asp Ile Val Leu Lys Trp Glu Leu
500 505 510

Gly Glu Gly Ala Phe Gly Lys Val Phe Leu Ala Glu Cys His Asn
515 520 525

Leu Leu Pro Glu Gln Asp Lys Met Leu Val Ala Val Lys Ala Leu
530 535 540

Lys Glu Ala Ser Glu Ser Ala Arg Gln Asp Phe Gln Arg Glu Ala
545 550 555

Glu Leu Leu Thr Met Leu Gln His Gln His Ile Val Arg Phe Phe
560 565 570

Gly Val Cys Thr Glu Gly Arg Pro Leu Leu Met Val Phe Glu Tyr
575 580 585

Met Arg His Gly Asp Leu Asn Arg Phe Leu Arg Ser His Gly Pro
590 595 600

Asp Ala Lys Leu Leu Ala Gly Gly Glu Asp Val Ala Pro Gly Pro
605 610 615

Leu Gly Leu Gly Gln Leu Leu Ala Val Ala Ser Gln Val Ala Ala
620 625 630

Gly Met Val Tyr Leu Ala Gly Leu His Phe Val His Arg Asp Leu
635 640 645

Ala Thr Arg Asn Cys Leu Val Gly Gln Gly Leu Val Val Lys Ile
650 655 660

Gly Asp Phe Gly Met Ser Arg Asp Ile Tyr Ser Thr Asp Tyr Tyr
665 670 675

Arg Val Gly Gly Arg Thr Met Leu Pro Ile Arg Trp Met Pro Pro
680 685 690

Glu Ser Ile Leu Tyr Arg Lys Phe Thr Thr Glu Ser Asp Val Trp
695 700 705

Ser Phe Gly Val Val Leu Trp Glu Ile Phe Thr Tyr Gly Lys Gln
710 715 720

Pro Trp Tyr Gln Leu Ser Asn Thr Glu Ala Ile Asp Cys Ile Thr
725 730 735

Gln Gly Arg Glu Leu Glu Arg Pro Arg Ala Cys Pro Pro Glu Val
740 745 750

Tyr Ala Ile Met Arg Gly Cys Trp Gln Arg Glu Pro Gln Gln Arg
755 760 765

His Ser Ile Lys Asp Val His Ala Arg Leu Gln Ala Leu Ala Gln
770 775 780

Ala Pro Pro Val Tyr Leu Asp Val Leu Gly
785 790

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TGYGAYATHA TGTGGYTNAA RAC 23

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGGATGCARY TNTGGCARCA RCA 23

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

YTCRTCYTTN CCRTAYTCRT T 21

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCYTCYTGRT ARTAYTCNAC GTG 23

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CACGTCAACA ACGGCAACTA CA 22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGAAGGATGA GAAACAGATT TCTGC 25

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CATCAATGGC CACTTCCTCA AGG 23

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGTGTTTCG TCCTTCTTCT CC 22

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GAGATGTGCC CGACCGGTTG TATC 24

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CACAGTGATA GGAGGTGTGG GA 22

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGATGTGGCT CCAGGCCCC 19

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGGCAACCCG CCCACGGAA 19

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ACGCCAGGCC AAGGGTGAG 19

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TAACCACTCC CAGCCCCTGG 20

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TTGGTGGCCT CCAGCGGCAG 20

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AATTCATGAC CACCAGCCAC CA 22

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCTCCTCGGG ACTGCGATGC 20

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATGTCGCCCT GGCCGAGGTG GCAT 24

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AAGCTCAACA GCCAGAACCT C 21

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CAGCTCTGTG AGGATCCAGC C 21

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCGACCGGTT TTATCAGTGA C 21

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ATGATCTTGG ACTCCCGCAG AGG 23

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTTGGCCAAG GCATCTCCGG T 21

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATGTGCAGCA CATTAAGAGG A 21

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TTATACACAG GCTTAAGCCA TCCA 24

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AGGAGGCATC CAGCGAATG 19

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Glu Ser Thr Asp Asn Phe Ile Leu Phe
1 5 9

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Leu Phe Asn Pro Ser Gly Asn Asp Phe Cys Ile Trp Cys Glu
1 5 10 14

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TCTCCTTCTC GCCGGTGG 18

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Ser Pro Ser Arg Arg Trp
1 5 6

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Phe Val Leu Phe His Lys Ile Pro Leu Asp Gly
1 5 10 11

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 84 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```
Trp Val Phe Ser Asn Ile Asp Asn His Gly Ile Leu Asn Leu Lys
 1               5                  10                  15

Asp Asn Arg Asp His Leu Val Pro Ser Thr His Tyr Ile Tyr Glu
            20                  25                  30

Glu Pro Glu Val Gln Ser Gly Glu Val Ser Tyr Pro Arg Ser His
            35                  40                  45

Gly Phe Arg Glu Ile Met Leu Asn Pro Ile Ser Leu Pro Gly His
            50                  55                  60

Ser Lys Pro Leu Asn His Gly Ile Tyr Val Glu Asp Val Asn Val
            65                  70                  75

Tyr Phe Ser Lys Gly Arg His Gly Phe
            80                  84
```

8

1